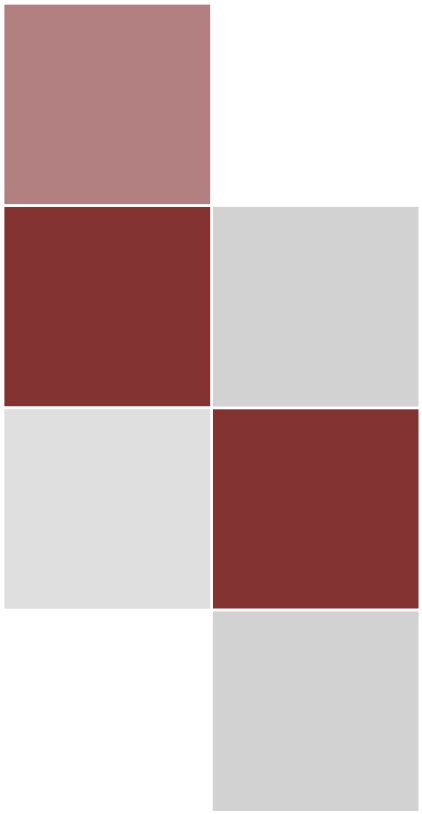
CGIAR

ILRI
INTERNATIONAL LIVESTOCK RESEARCH INSTITUTE



ILRI policy on research data management and sharing

March 2015

Citation: International Livestock Research Institute, 2015. ILRI policy on research data management and sharing. Nairobi: ILRI.

Ilri.org
better lives through livestock
ILRI is a member of the CGIAR Consortium

Box 30709, Nairobi 00100, Kenya
Phone : + 254 20 422 3000
Fax : +254 20 422 3001
Email: ILRI-Kenya@cgiar.org

Box 5689, Addis Ababa, Ethiopia
Phone: +251 11 617 2000
Fax : +251 11 617 2001
Email : ILRI-Ethiopia@cgiar.org

Contents

Version name	Version date	Who	Details
ILRI_RDM_Policy_Mar15	9 March 2015	Jane Poole	Approved IMC version; sent to BoT for approval in April 2015

Introduction

The International Livestock Research Institute (ILRI) works with partners worldwide to enhance the roles that livestock play in food security and poverty alleviation, principally in Africa and Asia. The outcomes of these research partnerships help smallholder farmers in developing countries keep their animals alive and productive, increase and sustain their livestock and farm productivity, find profitable markets for their animal products, and reduce the risk of livestock-related diseases.

ILRI leads the CGIAR Research Program (CRP) on Livestock and Fish (L&F), a component of the CRP on Agriculture for Nutrition and Health on the prevention and control of agriculture-associated diseases (A4NH), and contributes to six other CRPs.

ILRI's strategy (2013–2022) builds on the institute's established and unique global pro-poor livestock mandate, its evidence-based conviction that livestock have a vital role to play in enhancing food and nutritional security and reducing poverty in developing countries and its existing expertise, research agenda and partnerships (ILRI, 2013). The roles of the institute in its work with partners are articulated in three strategic objectives:

- Develop, test, adapt and promote science-based practices that—being sustainable and scalable—achieve better lives through livestock.
- Provide compelling scientific evidence in ways that persuade decision-makers—from farms to boardrooms and parliaments—that smarter policies and bigger livestock investments can deliver significant socio-economic, health and environmental dividends to both poor nations and households.
- Increase capacity among ILRI's key stakeholders and the institute itself so that they can make better use of livestock science and investments for better lives through livestock.

Data are ILRI's most important resources for research and decision making. The data generated by the institute are deemed to be international public goods (IPGs)[1] and should be accessible and usable not only to ILRI's partners but to the public.

This data management and sharing policy aims at supporting ILRI's strategic objectives, particularly objective two by storing, organizing and making available datasets generated in the course of research related activities for better evidence based decisions. It also aligns to the CGIAR open access and data management policy[2] and the CGIAR principles on the management of intellectual assets[3].

[1] International public goods (IPG's)are research outputs of knowledge and technology generated through strategic and applied research that are applicable and readily accessible internationally to address generic issues and challenges. Within the open access context, products are more likely to be IPG's if they are in searchable repositories (electronic); are globally available; are open and easily accessible.

[2] www.cgiar.org/open

[3] www.cgiar.org/resources/cgiar-intellectual-asset-management

CGIAR open access and data management policy

In November 2013, all 15 members of the CGIAR Consortium unanimously endorsed the Open Access and Data Management Policy designed to make final CGIAR Consortium information products – including publications, datasets, and audiovisual materials – Open Access ("OA").

Gains for ILRI

While the purpose of this policy is to enhance the role of data in supporting evidence-based decision-making, specific gains for ILRI and ILRI researchers are:

- Open Access ensures a wider dissemination of ILRI data thus improving transparency, participation, self-empowerment, innovation and recognition.
- Open Access data combined with high quality metadata also allows for straightforward and meaningful use of the data by both ILRI researchers and others, increasing the opportunities for influencing external decision-makers in the research for development community.
- Open Source tools provide alternative and free options that are accessible to a wide range of partners.
- With proper documentation Open Source software can be easily extended or integrated with other systems.
- Data stored in a repository combined with a good auditing process provides high levels of confidence in the quality of both the data and any derived products.
- Data residing in a server-based repository, backed-up and accessed by secured connections, removes the danger of data loss if individual computers crash.
- A well-defined data management process provides a clear plan for the treatment of ILRI data from collection to final storage with clear roles and responsibilities of people for the different elements.
- Sharing data in 'real-time' with partners and jointly validating and cleaning the data reduces the time between data collection and the production of the derived outputs.
- Connectivity and data interoperability allows third party organizations and systems to access ILRI data and exposes it to wider audiences.

Goal

The goal of this policy is to maximize the value of ILRI's data. This will be achieved when data are used to: facilitate research, deliver high quality research outputs and made available as international public goods for use by the global scientific community.

The purpose of this policy and accompanying guidance on implementation[4] is to provide both the centre's commitment to research data management & sharing and the framework for its practical implementation.

[4] ILRI research data management and sharing guidelines (www.ilri.org/open)

Objectives

To realise this goal, ILRI aims to meet FIVE objectives in research data management and sharing:

A. To ensure ILRI research data products support the development challenges in ILRI's strategy and those of ILRI focus CRPs;
B. To produce high quality research data products (well designed, collected, managed, verified and documented) – encouraging appropriate levels of standardization and harmonization so that the knowledge gained from the data is 'greater than the sum of its parts' whilst avoiding duplication;
C. To ensure ILRI research data are well protected and stored in suitable repositories;
D. To make ILRI research data available and accessible to all while respecting the rights of stakeholders in terms of confidentiality and intellectual property and data ownership;
E. To make research data products that are adaptable so that they can be easily transformed into actionable knowledge by ILRI and the global scientific community.

Scope

This policy applies to all data collected for the purposes of ILRI research, including the metadata that describe it and information derived from it. The policy applies to numerical and non-numerical data, and to data stored in both soft and hard copy form.

The majority of ILRI's institutional policies refer, in some way, to 'data' (e.g. ownership and sharing in ILRI Intellectual Property Rights policy, electronic storage in ILRI ICT policies). This policy attempts to minimise duplication of effort and overlaps by focussing specifically on issues relating to the management of data collected during research activities. Consequently, this policy should be followed in conjunction with other CGIAR and ILRI relevant policies on management of intellectual assets[5], research ethics[6], publishing guidelines, information asset classification and control and researcher good conduct.

This policy applies to all ILRI staff (including joint appointees), consultants, seconded staff, visiting scientists, postgraduate fellows and students involved in ILRI research, irrespective of their duty post. Staff members who leave ILRI are bound by this policy for 12 months after their departure.

To ensure research data management standards within joint projects with CGIAR and non-CGIAR partners the following guidance should be followed: if ILRI is responsible for project data management, then this policy applies; if a third party project partner is responsible, then the project contractual agreement must specify the principles which will guide the data management and sharing strategy. In the absence of such principles, then this policy will apply. These same conditions apply to other types of ILRI contract – e.g. CRP program participant and partner implementation agreements, memorandums of understanding, collaborative research agreements, teaming agreements etc.

[5] ILRI Management of Intellectual Assets Policy – https://cgspace.cgiar.org/handle/10568/34091
[6] ILRI Institutional Research and Ethics Committee (IREC) Operational Guidelines – http://ilri.org/ethicscommitteepoliciesandguidelines

Guiding principles

This policy is governed by key principles to ensure the achievement of ILRI's goal and objectives, details of these are given in Annex 1 and support on their implementation is provided in separate implementation guidelines:

- A positive research data management culture
- A well-defined research data management process
- Quality assurance standards
- Open Access and Open Source
- Ethics - data ownership & authorship, privacy & rights of stakeholders
- Documentation and metadata
- Suitable repositories, storage, preservation and security
- Interoperability, limited connectivity and translations
- Learning across research activities

Supporting environment

A supporting institutional environment will be necessary for the implementation of this policy. In particular, for ILRI's partner organisations who may have differing data management systems and capacities. This supporting environment includes:

1. Providing guidelines for managing and making data available in ways that are affordable while maintaining high quality standards.
2. Providing tools and systems for better and automated data management.
3. Supporting data entry tool alternatives depending on the budget and capacities of partners.
4. Sharing ideas across staff and partners and utilising tools and systems that build on previous experience.
5. Creating, maintaining and supporting portals for sharing and making data available.
6. Ensuring that performance assessment of staff[7] includes incentives, recognition and rewards for making research data more open and accessible.

Implementation

The implementation of this strategy is covered in a separate document[8]. The process of implementation addresses each principle of this strategy by providing guidance, tools and systems for putting these into practice.

Assessment, impact and review

This data management and sharing policy and its supporting documents and tools should be reviewed and updated as needed. Procedures and systems should be in place to monitor usage of ILRI's open access data resources. Efforts should also be made to link external and internal resources to ILRI's datasets as this adds greater value to our data and provides an indication of impact.

[7] staff (including joint appointees), consultants, seconded staff, visiting scientists, postgraduate fellows and students involved in ILRI research

[8] ILRI research data management and sharing guidelines (www.ilri.org/open)

Funding

ILRI projects (including CRPs) will adequately plan and budget for data management and open access. At an institute level funds will also be allocated to maintain ILRI data management infrastructure including repositories and online portals.

Acknowledgements

This policy refers to several documents already acknowledged in the footnotes. Other policies reviewed that provided insights and perspectives include: partner CGIAR centres and CRPs (CCAFS Data management strategy and associated documents, CIFOR Research Data Management Strategy and ICRAF Research Data Management Strategy); as well as others (PLOS Data Open-Access Policy, DFID Research Open and Enhanced Access Policy).

Annex 1: Details on guiding principles

A positive data management culture

People are at the core of this strategy. It must be made clear to ILRI staff (including joint appointees), consultants, seconded staff, visiting scientists, postgraduate fellows and students involved in ILRI research as well as partner organizations the advantages to them and to the institute of good data management, making data open access and their roles and responsibilities in the process. It is also critical to acknowledge the way people approached data management in the past, incorporate their views in the process, provide support and allow time for transition.

To enable this environment ILRI shall support the data management capacity development of staff, consultants, seconded staff, visiting scientists, postgraduate fellows, students and partners involved in the data management of ILRI projects (including CRP activities). Capacity development shall not be limited to the usage of systems and portals but include the change to, and maintenance of, a positive data management culture.

A well-defined research data management process

A clearly defined and documented data management process is needed for all ILRI projects (including CRP activities) even where different systems and tools for collecting data have been used[9]. However, the level of complexity and detail required will depend on the size of the research activity. The process should define staff roles and responsibilities as well as the strategy for ensuring the most efficient production of the data products. The actual process shall ensure high levels of quality and credibility while balancing the data management capacities of ILRI staff and partners.

Quality assurance standards

Explicit data validation procedures shall be used and documented throughout the research process; this includes proper logging of changes made to raw data as part of the cleaning process. The overall process as defined above also defines and ensures quality assurance.

The Centre has a zero tolerance policy with respect to data manipulation or falsification. All data sets shall have proper logging of changes including date, time and person committing any modifications to the raw data.

Open access and open source

The Centre views its research data as international public goods and is committed to their widespread diffusion and use to achieve the maximum possible access, scale, scope of impact and sharing of benefits to advantage the poor in developing countries.

Open access means the immediate, irrevocable, unrestricted and free online access by any user worldwide to information products, and unrestricted re-use of content (which could be restricted to non-commercial use and/or granted, subject to appropriate licenses), subject to proper attribution (CGIAR 2013)

[9] According to the 'CGIAR open access and data management implementation guidelines', ILRI should prepare open access and data management plans to ensure implementation of this policy. Such plans shall, in particular, outline a strategy for maximizing opportunities to make information products open access. http://library.cgiar.org

This policy considers that open data will serve ILRI's mission better. Key concerns here are to balance the rights of the researcher who generated the data (i.e. getting credit for their input) with the duty of the Centre (i.e. getting the most value from the data, irrespective of who generated them).

The Centre shall promptly and broadly disseminate through internet portals its research data to the public, including metadata and related research products; as per the CGIAR guidelines (4.2.4): *'Data (and any relevant data collection and analysis tools) shall, subject to any additional donor requirements, be deposited in a suitable repository and made Open Access as soon as possible and in any event within 12 months of completion of the data collection or appropriate project milestone, or within 6 months of publication of the information products underpinned by that data, whichever is sooner. Data deposited shall be prepared in a manner consistent with the aims of this Policy. Existing and future databases shall be made Open Access.'*

As per the conditions of some peer-reviewed journal publications and to improve scientific publications, consensus with scientific peers and public trust in the quality of ILRI research data, the Centre will provide institutional support to ensure that all necessary raw data will be made public to reproduce or replicate every scientific publication that is based on research data.

Open source - All tools used to process, visualize, store and make accessible data should also be Open Access and covered with a license to allow users to remix, tweak, and build upon ILRI's work. As different licenses may allow or restrict the use of derived tools for commercial use, it is necessary to evaluate the appropriate license depending on the potential of the tool.

Ethics - Data ownership and authorship, privacy and rights of stakeholders

Confidentiality of data on human subjects is respected in accordance with well-established international standards[10]. The details of how this will be achieved should be well-documented in agreements with partners and donors and details of how this is being achieved documented in the data management process for the activity.

Research data produced for ILRI staff (including joint appointees), consultants, seconded staff, visiting scientists, postgraduate fellows and students are defined as proprietary assets. Any Intellectual Property Rights associated with research data shall reside with the Centre, not the individual. Due recognition of the contributions of individuals shall be made both in the Open Access data and any publications associated with the data[11].

Documentation and metadata

Detailed information about the structure and type of dataset is necessary to produce meaningful data analysis and to ensure the data may be used by others; metadata includes information about the research activity that generated the data, relevant sampling or experimental design protocols, statistical analysis scripts and other related tools used to process the datasets and that provide meaning to the data. Derived information products such as scientific publications, policy briefs or books may be included as part of the metadata but in the form of links to such resources. Details of metadata requirements for ILRI data are provided in the Implementation document - "Minimum Metadata Standards".

[10] ILRI Management of Intellectual Assets Policy – https://cgspace.cgiar.org/handle/10568/34091 and IREC guidelines - http://ilri.org/ethicscommitteepoliciesandguidelines

[11] ILRI Publishing Guideline 4 - Authorship (www.ilri.org/open)

Suitable repositories, storage, preservation and security

A suitable repository should be able to store, securely preserve and allow concurrent access to datasets while keeping the integrity of the data and tracking of the changes made by users.

All data shall be stored in a suitable repository. The repository must have a backup policy and procedure in place.

Interoperability, limited connectivity and translations

Interoperability denotes the ability of diverse systems to communicate and work together (inter-operate). For this to happen, common protocols and standards must be stablished when publishing information.

Under the scope of this policy two levels of interoperability are necessary:

- Catalogue Interoperability: The ability to expose ILRI's catalogue of datasets to other systems
- Data Interoperability: The ability to expose the data itself of a particular dataset to other systems.

The ILRI dataset catalogue should be interoperable with other data and publications systems (e.g. data portals of partner organizations like UN, Worldbank and cataloguing systems like DataVerse, CGSpace). ILRI dataset should also be interlinked with other datasets and resources both internally and externally so they become more useful and better contextualized.

Efforts should be made to provide access to ILRI's datasets to stakeholders that have very limited internet connectivity. In addition, when data are collected using a different language than English, efforts should be made to provide the data in both English and the original language.

Learning across research activities

To get the most out of ILRI data the Centre and ILRI focus CRPs should learn across projects and research activities using cross-cutting indicators relating to the strategic development goals of the institute and CRP's and associated Monitoring, Evaluation and Learning (MEL) strategies. These may include indicators relating to food security, nutrition, poverty alleviation and gender equity, among others. Primary data collection in ILRI bilateral projects and CRPs should be designed to provide commonality in the generation of key indicators. Summaries of individual indicators can then be extracted from the datasets. It is the responsibility of the Centre to provide a data management infrastructure and support to researchers to achieve this[12].

[12] http://data.ilri.org/portal/, http://data.ilri.org/geoportal, http://data.ilri.org/tools